REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)

2. REPORT DATE
December 21, 1994

3. REPORT TYPE AND DATES COVERED
Final Report (3/91 - 10/94)

4. TITLE AND SUBTITLE
Object Oriented Segmentation of Images

5. FUNDING NUMBERS
DAAL03-91-G-0041

6. AUTHOR(S)
Jayant Shah

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Northeastern University
Boston, MA 02115

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 27790.6-MA

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

The central theme of this research project is recovery of object shapes from noisy images. The main mathematical techniques are energy functionals and gradient descent. The goal is to construct vertically integrated models capable of incorporating constraints imposed by various objectives such as noise suppression, boundary detection, shape description and object recognition. The difficulty with this approach is that these functionals are very hard to implement on account of their nonlinerity and the need to find discontinuous solutions. Therefore, the main emphasis of this research project has been on the formulation of approximate models. The key ingredient is representation of the discontinuity locus by a continuous variable which is then used to control processes at hand such as noise suppression or stereo matching. Specific formulations are developed for nonlinear smoothing of images, stereo matching which takes into account half-occlusions, recovery of surface shapes from shading data in the presence of noise and discontinuities and finally, curve evolution for locating and smoothing object boundaries.

14. SUBJECT TERMS
Image Processing, Shape Recovery

15. NUMBER OF PAGES
18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

OBJECT ORIENTED SEGMENTATION OF IMAGES

FINAL REPORT

Jayant Shah

December 21, 1994.

Accesion For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and / or Special
A-1	

U.S. ARMY RESEARCH OFFICE

GRANT NO. DAAL03-91-G-0041

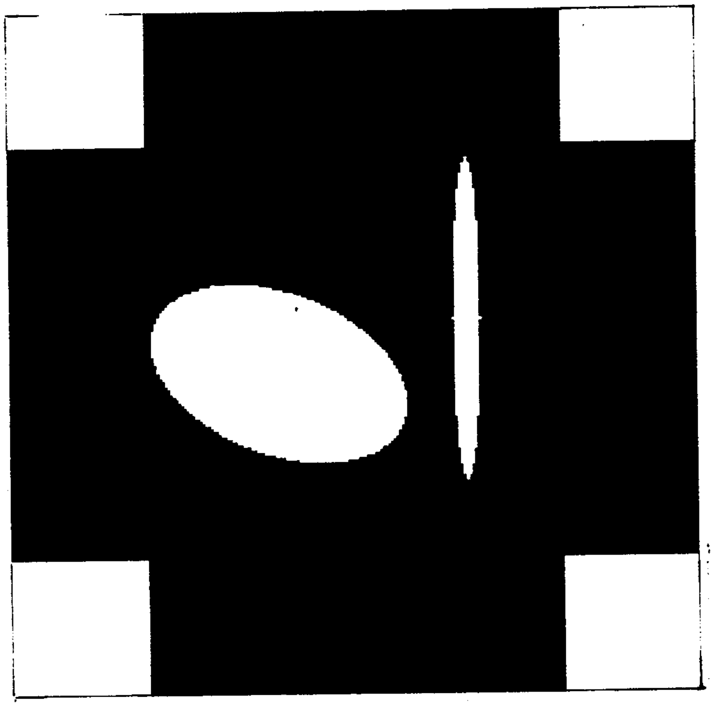
Northeastern University
Boston, MA 02115

APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

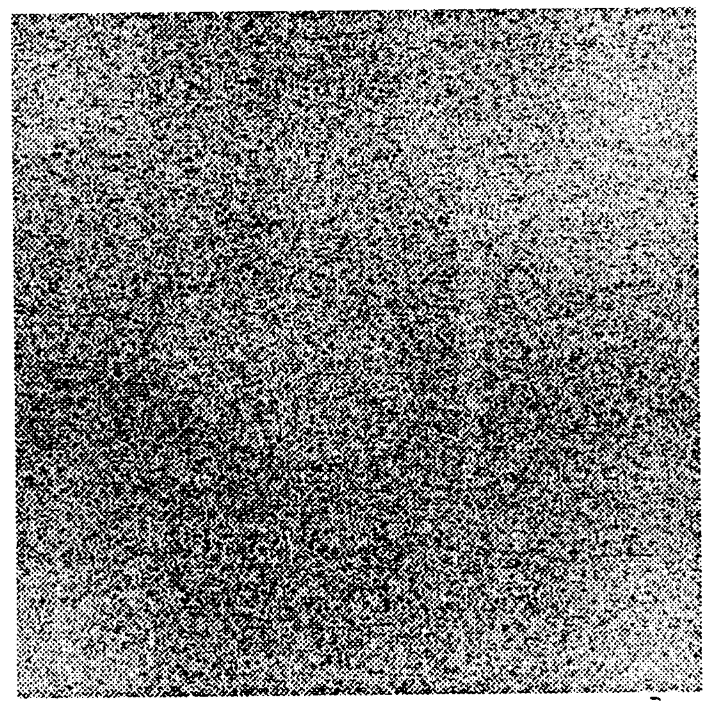
TABLE OF CONTENTS

1. INTRODUCTION

The central theme of this research project is recovery of object shapes from noisy images. Immediately, there is the question of what is a shape and how is it to be represented. Since answers to these questions have to be ultimately tailored to the uses one has in mind, one has to bring into consideration potential applications and with it, the question of practical algorithms for implementation of the theory. This research project is concerned with mainly two-dimensional shapes. Mathematically, a shape is the boundary of an object which is simply an open subset in the image domain, characterized in some way. In the real world, what is an object and what is just noise or clutter depends of course on what one is looking for. For example, Figure 1a shows a noiseless synthetic image. It is reasonable to assume that the objects in the figure are the four squares in the four corners and the two ellipses in the middle. Figure 1b shows a noisy version obtained from the image in Figure 1a by adding Gaussian noise. The signal-to-noise ratio (i.e. the ratio between the standard deviation of the image with noise removed and the standard deviation of the noise) is 1:4. The problem is to recover the original objects.



(a) NOISE-FREE IMAGE



(b) NOISY VERSION OF THE SAME IMAGE

FIGURE 1

The long term goal of this research project is shape analysis and object recognition. A common approach to the problem of shape representation and shape analysis is to assume that the shape boundary has already been found and the problem is to extract its salient features for subsequent use in object recognition. There are several approaches to this. One that is appropriate for characterizing shapes of objects subject to a group of transformations is to calculate some of the group invariants. The groups that are usually considered are subgroups of the group of projective transformations. This approach is not appropriate when homotopical or even topological alterations of shapes must be taken into account. Other approaches include finding differential invariants such as the points of maximum curvature or inflection and representation of shapes by shape skeletons such as the medial axis transforms [PBCFM]. Yet another approach is to describe a shape in terms of its parts such as necks and lobes. The crucial question from the point of view of this research project is how to recover these descriptions from noisy images in a robust way.

The main mathematical techniques for achieving the goals of this research project are energy functionals and gradient descent. The general approach is the construction of vertically integrated models which incorporate constraints imposed by various objectives such as noise suppression, boundary detection, shape description and object matching. The motivation for building integrated models comes from a long history of attempts to isolate, formulate and solve simpler problems in computer vision and the realization that no matter how sophisticated a technique is, there are inherent ambiguities which cannot be resolved in isolation. It is essential to incorporate contextual information in the model. Moreover, these methods are ideal for formulating models based on raw intensity images. This is in contrast to the earlier hierachical approaches in which features are extracted first and then these "feature maps" are processed for higher level tasks. The trouble is that the features cannot be identified with full accuracy without bringing in some higher level knowledge, and thus incorrect interpretations may be assigned during the construction of the feature map. By incorporating the raw image in the model, one can ensure that the input to higher level tasks such as object recognition reflects realistic ambiguities and noise, and thus ensure robustness of the solution. At the same time, integration provides further constraints which may help resolve ambiguities in the image. The disadvantage of such an approach is that as more and more realistic representations are incorporated in the energy functionals, they become more and more difficult to analyze and implement. Efforts to develop fast stochastic algorithms have so far not been very successful. An alternative approach is to construct approximations of these complex functionals, even deliberately weakening the coupling among their components and attempt to find stationary solutions by gradient descent.

2. SEGMENTATION PROBLEM

2.1 Segmentation Functional

Traditionally, the first step in image understanding is the detection of edges. The usual method is to smooth the image to reduce noise and then apply some kind of local edge operator such as the zero-crossings of the laplacian. Geman and Geman [GeGe] proposed a Bayesian formulation for simultaneous smoothing and boundary detection. At about the same time, Blake and Zisserman [BZ_1,BZ_2] proposed an analogous discrete model. An analytic version of these models is formulated in [MS_1] as follows:

$$E(u, B) = \frac{1}{\sigma^2} \int_R (u-g)^2 dx dy + \int_{R \setminus B} \|\nabla u\|^2 dx dy + \nu |B| \tag{1}$$

where

R is the image domain;
g is the feature intensity; $g : R \to \mathbf{R}$;
B is the union of segment boundaries; thus B is the segmenting curve;
$|B|$ is the length of B;
σ and ν are the weights. (σ may be thought of as the smoothing radius in $R \backslash B$.)

The task is to find u and B which minimize $E(u, B)$. Thus the segmentation problem is viewed as the problem of finding a piecewise smooth approximation of g with minimal amount of discontinuity. The weights σ and ν control the relative degrees to which smoothness of u and its discontinuity locus are emphasized. Smoothing is explicitly prevented from extending over B. An object in this formulation is characterized by having relatively uniform feature intensity. Many important theoretical results concerning existence and regularity have been obtained [Am,DMS,DCL,[MS_2],Ri,Sh_4,Wa]. In particular, in [MS_2] it is shown that the minimal length constraint implies that there cannot be any points where more than three objects meet and an object cannot have corners except at points where three objects meet. At such corners, the angle is always 120°. Thus the formulation implies a very special kind of shape representation and it was one of the goals of this research project to investigate ways in which formulation (1) may be altered or augmented to represent shapes more realistically.

2.2 Dirichlet Version

A simple way to preserve the actual shape singularities is to impose Dirichlet boundary conditions on u in functional (1), namely, $u^{\pm} = g^{\pm}$ along B where the superscript $\pm$ denote the values on the two sides of B. Results concerning the existence and approximations in this case are proved in [Sh_6]. It is shown that singularities of u are then a subset of those of g; thus distortion of corners and junctions is prevented. The limitation of this modified model is that if the actual boundaries are noisy (i.e. locally the graph of a noisy function), they will not be smoothed. Moreover, if the boundaries in the image are blurred as is likely in practical situations so that g is continuous, then B would be empty. Hence the formulation must always be used with an explicit representation for blurring. Such a representation based on Ambrosio-Tortorelli approximation (described below) is formulated and analyzed in [Sh_6]. Note that blurring also provides smoothing of noisy boundaries. Recovery of unblurred boundaries from their blurred version is treated as a separate problem.

2.3 Extensions

An obvious extension for enriching shape representation in functional (1) is to incorporate information about the curvature of the boundary. One such extension is proposed in [NM] where $|B|$ is replaced by

$$\int_B \left(\nu_0 + \nu_1 \kappa^2\right) ds \tag{2}$$

where κ denotes the curvature. Very few theoretical results concerning the new functional have been obtained. Results for simple closed curves based on elastica for functional (2) under the assumption that $v_0 = 0$ and that the length of the curve is fixed have been obtained in [M,W]. In [BDP], lower semi-continuity of functional (2) is analyzed and certain pathologies that arise are exhibited. For instance, even for a simple heart shape, minimizing sequences produce regions of zero width in the limit. Practical difficulties in implementing such a functional are severe and as yet, no one has implemented it. Allowing B to be only piecewise continuous or introduction of more shape information such as illusory contours or 3D interpretations makes the functional even more intractable. In view of these difficulties which increase as the functionals become more complex, it was considered very important to devote a major part of this research effort to the construction of approximations which are a compromise between formulations expressed by a single functional on one hand and the earlier bottom-up hierarchical formulations on the other.

2.4 Elliptic Approximation

What makes the task of minimization of functional (1) very difficult is the presence of the one-dimensional segmenting curve. Ambrosio and Tortorelli [AT_1,AT_2] proposed an elliptic approximation of functional (1) so that gradient descent may be applied. Their method is as follows. Let

$$\Lambda_\rho(v) = \frac{1}{2}\int_R \left\{\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right\} dx\,dy \tag{3}$$

For a fixed curve B, if v_ρ minimizes $\Lambda_\rho(v)$ with boundary condition $v = 1$ on B, then as $\rho \to 0$, v_ρ obviously tends to zero everywhere except on B where it equals one. The key observation is that $\Lambda_\rho(v) \to |B|$ as $\rho \to 0$. Values of v_ρ range between 0 and 1 and thus may be interpreted as the probability for the presence of a boundary point. Alternatively, v_ρ may be viewed as blurring of B with ρ as the nominal blurring radius. The method of Ambrosio and Tortorelli is to replace the term $|B|$ in functional (1) by $\Lambda_\rho(v)$. B is now spread out over all of R and we have to modify the integral in functional (1) as well since we no longer have the set $R\backslash B$. A simple choice is to

$$\text{replace} \int_{R\backslash B} \|\nabla u\|^2 dx\,dy \text{ by } \int (1-v)^2\|\nabla u\|^2 dx\,dy \tag{4}$$

The final form of the approximate functional is

$$E_\rho(u,v) = \int_R \left\{\frac{1}{\sigma^2}(u-g)^2 + (1-v)^2\|\nabla u\|^2 + \frac{\nu}{2}\left(\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right)\right\} dx\,dy \tag{5}$$

Ambrosio and Tortorelli prove that $E_\rho(u,v)$ converges to $E(u,B)$ in the sense of Γ–convergence.

The corresponding approximate functional when dirichlet boundary conditions are imposed was derived and it is as follows [Sh_6]:

$$\begin{aligned} E_{M,\rho}(u,v) = &\int_R \left[\frac{1}{\sigma^2}(u-g)^2 + (1-v)^2\|\nabla u\|^2\right] dx\,dy \\ &+ \int_R \left[\left\{\frac{\nu}{2} + M(u-g)^2\right\}\left\{\left(\rho\|\nabla v\|^2 + \frac{v^2}{\rho}\right)\right\}\right] dx\,dy \end{aligned} \tag{6}$$

where M is of order $O|\log \rho|$.

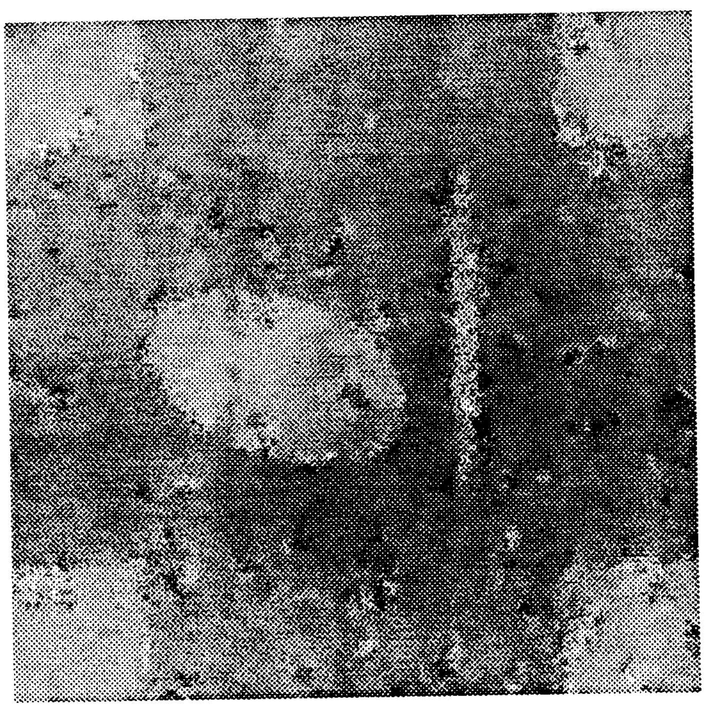
It should be noted that an approximation theory for functional (1) when terms containing curvature and singularities of B are introduced has yet not been found.

Stationary solutions for functionals (5) and (6) may be easily found by applying gradient descent. Thus, applying gradient descent to (5), we get the following coupled diffusion equations:

$$\begin{aligned} \text{Smoothing: } & \frac{\partial u}{\partial t} = \nabla \cdot (1-v)^2 \nabla u - \frac{1}{\sigma^2}(u-g); \quad \frac{\partial u}{\partial n}|_{\partial R} = 0 \\ \text{Boundary detection: } & \frac{\partial v}{\partial t} = \nabla \cdot \nabla v - \frac{v}{\rho^2} + \frac{2}{\nu\rho}(1-v)\|\nabla u\|^2; \quad \frac{\partial v}{\partial n}|_{\partial R} = 0 \end{aligned} \tag{7}$$

where ∂R denotes the boundary of R and n denotes the direction normal to ∂R.

The results of Ambrosio-Tortorelli diffusion (7) when applied to the example shown in Figure 1b with $\sigma = \rho = 8$ pixels are shown in Figure 2. The image in Figure 1b was represented on a 256×256 square lattice. The thin ellipse has the maximum width of 9 pixels and the distance between the two ellipses is 18 pixels. Note that long thin objects are very difficult to locate in noisy images because the smoothing radius is governed by the noise characteristics and thus may be much larger than the width of the object. Figure 2a shows the smoothed image u and Figure 2b depicts the edge strength function v. The lighter the area in the figure, the higher the value of v.



(a) SMOOTHED VERSION u

(b) EDGE-STRENGTH FUNCTION v

FIGURE 2

3. SHAPE RECOVERY

The use of approximate segmentation functionals discussed above for constructing practical algorithms leads to a new difficulty, namely, the difficulty of recovering the actual boundaries from the edge-strength function v. It is apparent from Figure 2b that thresholding of v will not produce satisfactory representation of the boundaries. The trouble is that due to differing levels of contrast along the boundaries and the surrounding noise, values of v along the boundary are not constant. That is, the level curves of v are only approximately parallel to the object boundaries; in fact, they might even be perpendicular to the object boundary in places. Consequently, global thresholding of v produces representation of the boundaries by narrow strips of varying width which may not completely enclose the object if the contrast becomes sufficiently weak.

3.1 Adaptive Thresholding

Schemes for adaptive local thresholding of v are discussed in [Sh3]. The basic idea is that the local threshold for v should depend on the average value v in the neighborhood. What this means is that we should look at the laplacian of a smoothed version of v. Ambrosio-Tortorelli equations are applied to v to produce such a smoothing. This approach to local thresholding of v does produce improvement, but such further processing was found to produce further displacement of the boundary from its actual location, indicating the need for a better method for recovering the actual boundary from v.

3.2 Shape Recovery by Curve Evolution

The basic idea is borrowed from the work of Kass, Witkin and Terzopoulos [KWT] on SNAKES. In their framework, one introduces a simple closed curve in the image and lets it deform under forces which are determined by distance from the object boundary and by a smoothness constaint. In the framework presented here, the analogous idea [Sh7] is to let an initial curve evolve so as to (locally) maximize the edge-strength function v along the curve in some sense. Let Γ be a simple closed curve in R. In order to move Γ to where the image intensity gradient and hence v are high, we look for the stationary points of the functional

$$L = \int_{\Gamma} (1 - v)^{\alpha} d\gamma \tag{8}$$

where γ denotes the arc-length along Γ. The evolution equation for Γ is derived by applying gradient descent to L. Let $C(p,t) : I \times [0,\infty) \rightarrow R$ be the evolving family of curves where I is the unit interval and t denotes time. We require that $C(0,t) = C(1,t)$ for all values of t and that the image of $C(p,0)$ in R coincides with Γ. Then the evolution of Γ is governed by the equation

$$\frac{\partial C}{\partial t} = [\alpha \nabla v \cdot N - (1 - v)\kappa] N \tag{9}$$

where N is the outward normal and κ is the curvature which is defined such that it is positive when Γ is a circle. Thus the points on the evolving curve move in the direction of the normal with velocity which has two components: the component depending on curvature which imposes a smoothness constraint and advection induced by v. The curve is pulled towards the object boundary by the force field ∇v induced by v. The exponent α in the expression for L serves as a weighting factor. The higher the value of α, the weaker the smoothing constraint.

The implementation of curve evolution is no longer a straightforward matter of using finite differences. Moving the points on an evolving curve directly by discretizing the curve leads to many difficulties. For example, the chosen points might bunch up causing numerical instabilities. The curve may also undergo topological changes. An alternate method is the one proposed by Osher and Sethian [OS]. In their approach, the initial curve is embedded in a surface as a level curve and then evolution is applied to the surface so that all of its level curves evolve simultaneously. Assume that Γ is embedded in a surface $f_0 : R - \mathbf{R}$ as a level curve. Let $f(t, x, y)$ denote the evolving surface such that $f(0, x, y) = f_0(x, y)$. Then, in order to evolve all the level curves of f_0 simultaneously, we consider the functional

$$G(f) = \int_{-\infty}^{\infty} \int_{\Gamma_c} (1 - v)^{\alpha} d\gamma_c dc \tag{10}$$

where $\Gamma_c = \{(x, y) | f(t, x, y) = c\}$. But by the coarea formula,

$$G(f) = \int \int_{R} (1 - v)^{\alpha} ||\nabla f|| dx dy \tag{11}$$

By calculating the first variation of the last functional, we get the gradient descent equation as

$$\begin{aligned} \frac{\partial f}{\partial t} &= -\alpha \nabla v \cdot \nabla f + (1 - v)||\nabla f|| \nabla \cdot \left(\frac{\nabla f}{||\nabla f||} \right) \\ &= -\alpha \nabla v \cdot \nabla f + (1 - v) \frac{f_y^2 f_{xx} - 2 f_x f_y f_{xy} + f_x^2 f_{yy}}{||\nabla f||^2} \end{aligned} \tag{12}$$

The boundary conditions are:

$$\begin{aligned} \frac{\partial f}{\partial n} &= 0 \text{ along the boundary of } R \\ f &= f_0 \text{ at } t = 0 \end{aligned} \tag{13}$$

For numerical implementation of equation (11), the usual finite-difference schemes (say, central differences) are exactly the wrong thing to apply and must never be used. The main point of Osher and Sethian is that since we expect the evolving surface to develop discontinuities or 'shocks' where the object boundaries are, the directions of the finite differences must always be away from the shock; a finite-difference must never straddle a shock. Once this principle is incorporated in the numerical scheme, nonlinear diffusion (11) behaves very robustly.

An important question now is: How to specify the intial curve Γ? An automatic specification of the initial curve is a difficult problem because its solution implies that the object boundaries have already been found, at least approximately. The strategy used in [Sh$_7$] for specifying the initial curve is based on the following considerations. If we set $v = 1$ and $f_0 = g$ in equation (11), then it reduces to the diffusion equation proposed in [ALM] for smoothing images. In [ALM], diffusion is allowed to occur only in the direction of the level curves, the expectation being that in this way, the object boundaries will be smoothed without being blurred. The difficulty with this approach is of course that the object boundaries in general are not level curves in an image. Traditional use of the zero-crossings of the laplacian of smoothed images to detect edges may be thought of as an attempt to remedy this situation by representing object boundaries by the level curves of the smoothed laplacian instead of the level curves of the image itself. Hence, the strategy in [Sh_7] is to use the zero-crossings of the laplacian as the initial approximation Γ for the object boundaries and choose f_0 equal to the laplacian of a smoothed version of the image. Now, if the raw image g is very noisy, the laplacian of the smoothed image u obtained from Ambrosio-Tortorelli diffusion will also be very noisy. To see this, consider the behavior of u in the limit as $\rho \rightarrow 0$. In the limit, u is a stationary solution of the original problem (1) and satisfies the differential equation $\nabla^2 u = (u-g)/\sigma^2$, indicating that the laplacian will be as noisy as the original image. It will have many saddle points, indicated by many self-intersections of its zero-crossings. The theory of curve evolution is based on the assumption that Γ is a simple closed curve, an assumption which is violated at the saddle points of f. If the laplacian is too noisy, the evolution is dominated by what happens at its saddle points and becomes unpredictable. This is because at a saddle point, the first term on the right hand side of the evolution equation (11) vanishes and both the numerator and the denominator vanish in the second term. Hence, the behaviour of the second term becomes very sensitive to noise near a saddle point. In order to demonstrate the feasibility of this approach, an ad hoc solution to this problem has been adopted in [Sh_7]. We smooth u further by non-uniform smoothing as described below until the zero-crossings of the laplacian have relatively few self-intesections. Recall that we can implement uniform Gaussian smoothing of u by applying gradient descent to the functional

$$\int\int_R \|\nabla w\|^2 dxdy \tag{14}$$

setting $w = u$ at $t = 0$. In order to reduce the displacement of the significant boundaries (indicated by high values of v) due to smoothing, we replace the above functional by the functional

$$\int\int_R (1-v)^\alpha \|\nabla w\|^2 dxdy \tag{15}$$

The corresponding gradient descent equation is

$$\begin{aligned} \frac{\partial w}{\partial t} &= -\alpha \nabla v \cdot \nabla w + (1-v)\nabla^2 w \\ \frac{\partial w}{\partial n} &= 0 \text{ along the boundary of } R \\ w &= u \text{ at } t = 0 \end{aligned} \tag{16}$$

The diffusion by equation (15) is stopped by the user when the level curves of $\nabla^2 w$ have relatively few self-intesections. Then we set $f_0 = \nabla^2 w$ and let it evolve according to equation (11). The zero-crossings of the evolving f are taken as the successive refinements of the object boundaries.

Figures 3 and 4 show the results of evolution when applied to the example shown in Figure 1b. The laplacian of u was so noisy that *almost all the pixels* were zero-crossings of the laplacian. Hence, u was further smoothed by nonuniform diffusion (15) and then evolution by equation (11) was applied to its laplacian. Figure 3 shows the zero-crossings of the evolving laplacian at $t = 0, 40, 80, 160, 320, 640, 1280$. Thus, the zero-crossings at $t = 0$ are the zero-crossings of the laplacian of u smoothed by equation (15). The example illustrates that the boundaries found by this method are at least metastable, that is, they persist for a long time. Figure 4 shows the superposition of the zero–crossings smoothed by evolution ($t = 1280$) on the noisefree image (Figure 4a) as well as on the smoothed image u (Figure 4b). Smoothed image u was used for superposition rather than g because the evolution is governed by v which, in turn, is determined by the boundaries in u. Note the accuracy of placement of the smoothed zero-crossings on the noiseless image, including the thin ellipse. The corners are also fairly well represented. The boundary deviates in places from the boundary in the noiseless image when it follows some accidental feature introduced by the noise. (It is possible to discern these features by close inspection of Figure 2a.) It is interesting that although the four corner squares are identical in the noiseless image, the final boundaries are different in all four cases because the accidental noise features are different. The worst deviation occurs in the case of the lower right corner square. The very thin ends of the thin ellipse are lost because the boundary representation v is too coarse ($\rho = 8$ pixels). Some portions towards the ends of the thin ellipse are lost because they are obscured by noise as one can see in Figure 2a.

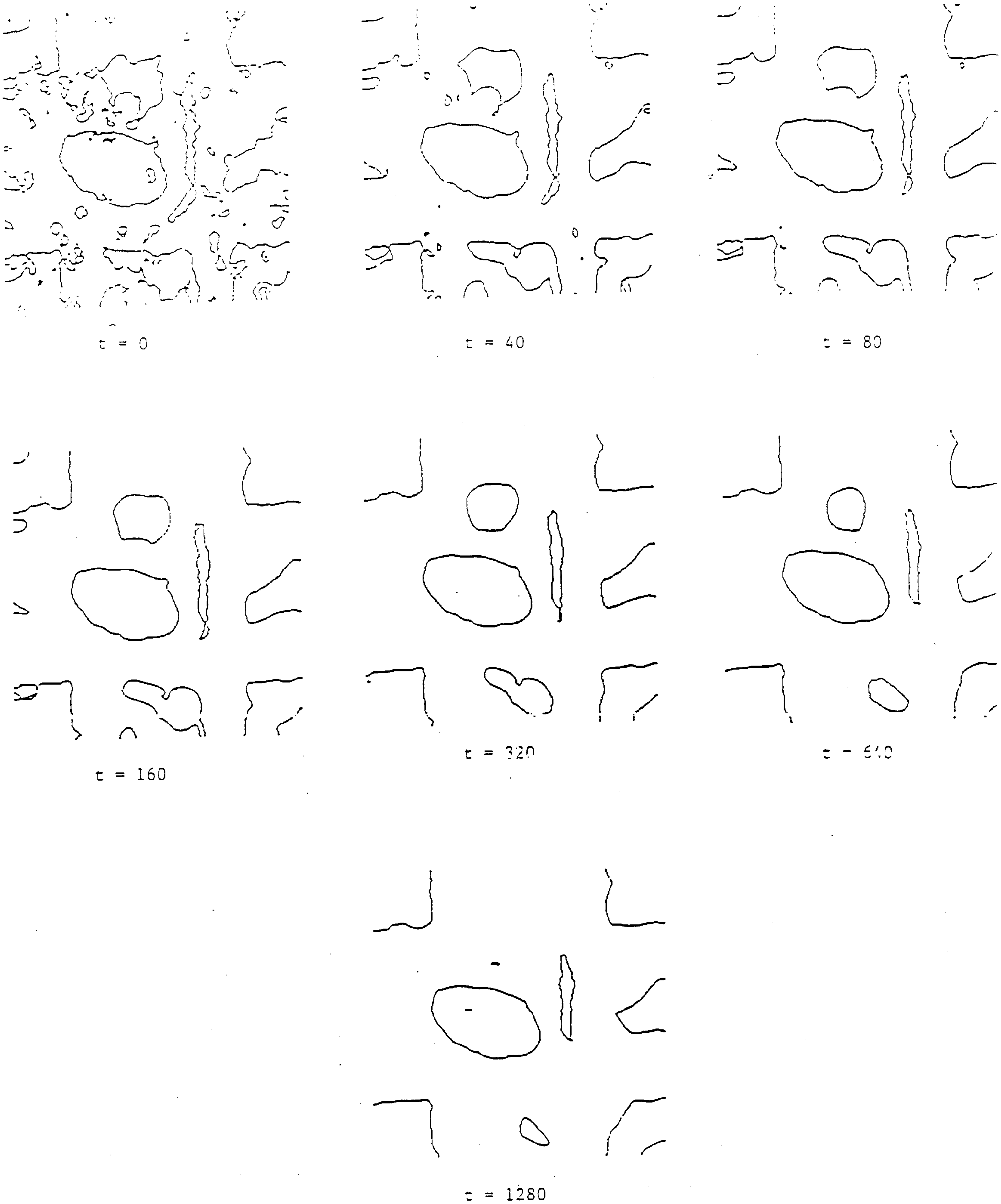
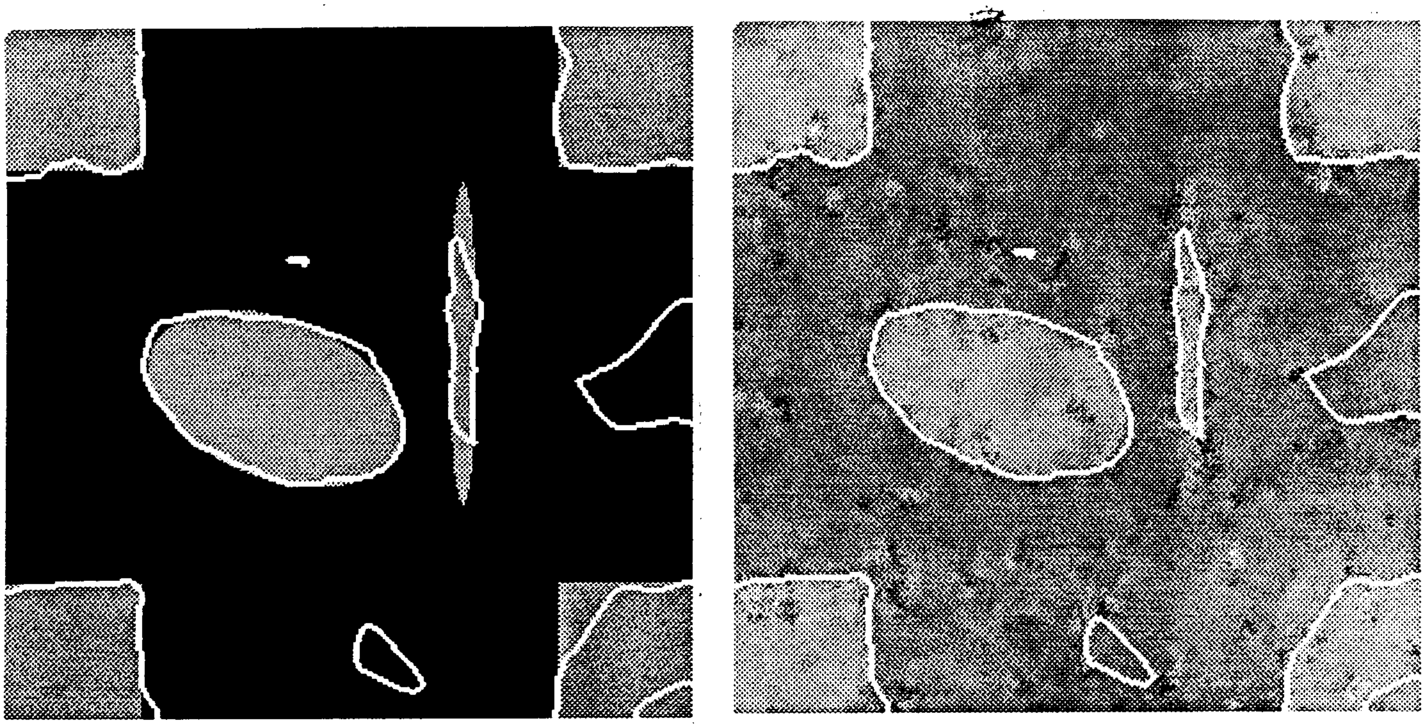


FIGURE 3 EVOLUTION OF ZERO_CROSSINGS



(a) SUPERPOSITION ON NOISE-FREE IMAGE (b) SUPERPOSITION ON SMOOTHED IMAGE u

FIGURE 4

4. DIFFUSION SYSTEMS

Use of curve evolution for shape recovery as described above exemplifies an approach for getting around the difficulties in finding the minimizing solutions for energy functionals. The basic ingredient was to incorporate the edge-strength function v in the algorithms designed to solve the next level of tasks in computer vision. If these tasks are also governed by evolution equations, then the idea is to let this evolution be controlled by v. In fully integrated models, one would try to construct a coupled system of diffusion equations in which each equation controls one particular process or a task like smoothing or edge detection or matching and outputs of these equations provide controls for each other's evolution. The diffusion system (7) of Ambrosio and Tortorelli may be viewed as a system of this type. Other examples are described below. The difficulty in designing nonlinear systems of diffusion equations for problems in computer vision is that such systems are one of the most difficult mathematical objects to analyze and no general mathematical principles are available to guide the design.

4.1 Stereo

In stereo vision, the basic problem is the correspondence problem, that is, the problem of matching the two views seen by the two eyes (or cameras) and thus computing the disparity between the two images. A major complication is the presence of discontinuities. Due to sudden changes in the distance of the objects from the eyes, there is a sudden jump in the disparity.

Another consequence of this is the phenomenon of half-occlusions which are the areas in the scene seen by one eye and not the other. Thus it is necessary to prevent matching of the two images in areas of half-occlusions. A coupled diffusion system of 4 equations is formulated in [Sh5] as follows:

$$\begin{aligned}
\frac{\partial d_\ell}{\partial t} &= \nabla \cdot (1-w_\ell)^2 \nabla d_\ell + \frac{1}{\sigma^2}(I_\ell - I_r \circ f_\ell)\left(\frac{\partial I_r}{\partial x_r} \circ f_\ell\right)(1 - w_r \circ f_r^{-1})^2 \\
\frac{\partial w_\ell}{\partial t} &= \rho \nabla \cdot \nabla w_\ell - \frac{w_\ell}{\rho} + \frac{2}{\nu}(1-w_\ell)\|\nabla f_\ell\|^2 \\
\frac{\partial d_r}{\partial t} &= \nabla \cdot (1-w_r)^2 \nabla d_r + \frac{1}{\sigma^2}(I_r - I_\ell \circ f_r)\left(\frac{\partial I_\ell}{\partial x_\ell} \circ f_r\right)(1 - w_\ell \circ f_\ell^{-1})^2 \\
\frac{\partial w_r}{\partial t} &= \rho \nabla \cdot \nabla w_r - \frac{w_r}{\rho} + \frac{2}{\nu}(1-w_r)\|\nabla f_r\|^2
\end{aligned} \tag{17}$$

where

d_ℓ, d_r are the disparities as seen from the left and right eyes respectively;
I_ℓ, I_r are the left and right images;
x_ℓ, x_r are the coordinates in the left and right images parallel to the epipolar lines;
$f_\ell = x_\ell + d_\ell, f_r = x_r + d_r$;
w_ℓ, w_r represent the discontinuity locus of disparities d_ℓ, d_r respectively.

The first and third equations are the stereo matching processes from the left and right eyes respectively, while the second and the fourth equations are discontinuity detection processes. The formulation is illustrated by examples in [Sh5]

4.2 Shape-from-Shading

The problem here is to solve the irradiance equation to recover the surface shape from the shading data. As usual, one has to account for noise and also discontinuities in surface shape in the form of creases. In [SHG], a formulation is derived for the shape-from-shading problem, taking into account noise and discontinuities, and also allowing fusion with data from other sources such as range images. The diffusion system is given as follows:

$$\begin{aligned}
&\frac{\partial f}{\partial t} = \nabla \cdot V + \frac{1}{\lambda^4}(d - f) \\
&\text{where vector } V = \frac{\alpha^2}{\lambda^4}(R - I)\nabla_{f_x,f_y} R - \left[\nabla \cdot (1-s)^2 \nabla f_x, \nabla \cdot (1-s)^2 \nabla f_y\right] \\
&\frac{\partial s}{\partial t} = \nabla \cdot \nabla s - \frac{s}{\rho^2} + \frac{2\lambda^4}{\nu\rho}\left(\|f_x\|^2 + \|f_y\|^2\right)(1-s)
\end{aligned} \tag{18}$$

Here,

f represents the surface to be recovered;
d is the range data;
I is the intensity image;
R is the reflectance function defined in terms of the derivatives f_x, f_y;
∇_{f_x,f_y} is the gradient operator with respect to variables f_x, f_y;
s represents the locus of creases in the surface f.

The first equation fits the surface to the range data and to the intensity image though the reflectance map. The surface is regularized by a second order smoothing constraint. The last equation detects discontinuities in the gradient of f and restricts smoothing of f wherever creases occur.

4.3 Nonlinear Smoothing

Alternate systems of diffusion equations, coupling a smoothing process and a boundary detection process in a manner analogous to the Ambrosio-Tortorelli system (7) are described in [Sh_3] and [PPGO].

4.4 Optical Flow

The framework of coupled diffusion systems developed in this research project has been used by Proesmans et al [PGO] to construct a system of six diffusion equations for calculating motions of objects moving at different speeds and possibly occluding each other.

BIBLIOGRAPHY

[ALM] L. Alvarez, P.L. Lions and J.M. Morel: "Image Selective Smoothing and Edge Detection by Nonlinear Diffusion II", SIAM J.Num.Anal, (June,1992).

[AT_1] L. Ambrosio and V.M. Tortorelli: "Approximation of Functionals Depending on Jumps by Elliptic Functionals via Γ-convergence", Preprint, Scoula Normale Superiore, Pisa, (1988)

[AT_2] L. Ambrosio and V.M. Tortorelli: "On the Approximation of Functionals depending on Jumps by Quadratic, Elliptic Functionals", Boll. Un. Mat. Ital. (1992).

[BDP] G. Bellettini, G. Dal Maso and M. Paolini: "Semicontinuity and Relaxation Properties of a Curvature depending Functional in 2d", SISSA Report, 17/92/MA, (Feb. 1992).

[BZ_1] A. Blake and A. Zisserman: "Using weak continuity constraints", Report CSR-186-85, Dept. of Comp. Sci., Edinburgh Univ. (1985)

[BZ_2] A. Blake and A. Zisserman: Visual Reconstruction, M.I.T. Press (1987)

[DMS] G. Dal Maso, J.M. Morel and S. Solimini: "A Variational Method in Image Segmentation: Existence and Approximation Results", Preprint, CEREMADE, Universite Paris - Dauphine, Paris (1988)

[DCL] E. De Giorgi, M. Carriero and A. Leaci: "Existence Theorem for a Minimum Problem with Free Discontinuity Set", Preprint, University of Lecce, Italy (1988)

[GeGe] S. Geman and D. Geman: "Stochastic relaxation, Gibbs' distributions, and the Bayesian restoration of images", IEEE Trans., PAMI 6, pp. 721-741 (1984)

[KWT] M. Kass, A. Witkin and D. Terzopoulos: "Snakes: Active Contour Models", First International Conf. on Computer Vision, (1987).

[M] D. Mumford: "Elastica and Computer Vision", Preprint.

[MS_1] D. Mumford and J. Shah: "Boundary detection by minimizing functionals, I", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco (1985)

[MS_2] D. Mumford and J. Shah: "Optimal approximations by piecewise smooth functions and associated variational problems", Comm. on Pure and Appl. Math., v. XLII, n.5, pp.577-684 (July, 1989)

[NM] M. Nitzberg and D. Mumford: "The 2.1D Sketch", Third Int'l Conf. on Comp. Vision, (December, 1990).

[OS] S. Osher and J. Sethian: "Fronts Propagating with Curvature Dependent Speed: Algorithms based on the Hamilton-Jacobi Formulation", J. Comp. Physics, 79, (1988).

[PBCFM] S.M. Pizer, C.A. Burbeck, J.M. Coggins, D.S. Fritsch, B.S. Morse: "Object Shape before Boundary Shape: Scale-Space Medial Axes", J. of Math. Imaging and Vision, (1993).

[PPGO] M. Proesmans, E.J. Pauwels, L.J. Van Gool and A. Oosterlinck: "Image Enhancement using Non-Linear Diffusion", Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1993).

[PGO] M. Proesmans, L.J. Van Gool and A. Oosterlinck: "Determination of Optical Flow and its Discontinuities using Non-linear Diffusion", Preprint, ESAT-MI2, Katholieke Univ. Leuven, Leuven, Belgium, ((1993).

[Ri] T. Richardson: Ph.D. Thesis, Department of Electrical Engineering and Computer Science, MIT (1990)

[Sh_1] J. Shah: "Parameter Estimation, Multiscale Representation and Algorithms for Energy-Minimizing Segmentations", Tenth International Conference on Pattern Recognition, (June, 1990)

[Sh_2] J. Shah: "Segmentation by Nonlinear Diffusion", Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1991).

[Sh_3] J. Shah: "Segmentation by Nonlinear Diffusion, II", Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1992).

[Sh_4] J. Shah: "Properties of Energy-Minimizing Segmentations", SIAM J. on Control and Optim. v.30, no.1, 99–111, (1992).

[Sh_5] J. Shah: "A Nonlinear Diffusion Model for Discontinuous Disparity and Half-Occlusions in Stereo", Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1993).

[Sh_6] J. Shah: "Piecewise Smooth Approximations of Functions", Calculus of Variations and Partial Differential Equations, 2, pp.315-328, (1994).

[Sh_7] J. Shah: "Recovery of Shapes by Evolution of Zero-Crossings", submitted for presentation at the Fifth International Conference on Computer Vision, (1995).

[SHG] J. Shah: "Role of Fusion and Regularization in the Shape-from-Shading Problem", submitted to Proc. IEEE Conf. on Comp. Vision and Pattern Recognition, (June, 1994).

[SK] K. Siddiqi and B.B. Kimia: "Parts of Visual Form: Computational Aspects", IEEE Conf. on Vision and Pattern Recognition, (1993).

[STK] K. Siddiqi, K. Tresness and B.B. Kimia: "Parts of Visual Form: Ecological and Psychological Aspects", Technical Report LEMS-104, Div. of Engineering, Brown University, (Feb. 1994).

[Wa] Y. Wang: Ph.D. Thesis, Dept. of Math, Harvard Univ, (1989).

[W] Y. Wen: "L^2 Flow of Curve Straightening in the Plane", Duke Math. J. v.70, n.3, pp.683-398, (June, 1993).

PUBLICATIONS

1. *Recovery of Shapes by Evolution of Zero-Crossings*, Submitted to International Conference on Computer Vision, to be held in May, 1995.
2. *Recovery of Shapes of Surfaces with Discontinuities by fusing Shading and Range Data within a Variational Framework*, (with H. Pien and J. Gauch), Submitted to IEEE Proceedings on Image Processing.
3. *Piecewise Smooth Approximations of Functions:* Calculus of Variations and Partial Differential Equations, v.2, pp.315-328 (1994).
4. *A Nonlinear Diffusion Model for Discontinuous Disparity and Half-Occlusions in Stereo:* IEEE Conference on Computer Vision and Pattern Recognition, New York, N.Y., June, 1993.
5. *Segmentation by Nonlinear Diffusion, II:* Proc. of the IEEE Conference on Computer Vision and Pattern Recognition,Champaign, IL, June, 1992.
6. *Segmentation by Nonlinear Diffusion* Proc. of the IEEE Conference on Computer Vision and Pattern Recognition,Lahaina, Hawaii, June, 1991.

SCIENTIFIC PERSONNEL: **Jayant Shah**